有効な配列空間を探索するタンパク質の進化分子工学の開発研究

はっとり　しゅうへい
氏　　名　服 部　修 平
研究科，専攻の名称　東北大学大学院工学研究科（博士課程）バイオ工学専攻
学 位 論 文 題 目　有効な配列空間を探索するタンパク質の進化分子工学の開発研究
論 文 審 査 委 員　主査　東北大学教授　梅津 光央　東北大学教授　中山 亨
東北大学教授　魚住 信之

論 文 内 容 要 約

First section: Introduction

Proteins were constituted by 20 kinds of amino acids and function based on amino acid sequences. Therefore, the structure and the function of proteins were not predicted from the amino acid sequence, because the possible sequence space for amino acids was massiveness not to analyze exhaustively. Thus, it suggested that directed evolution, which was the method that the objective variant was acquired by screened the library mutated to the scaffold protein determined from the protein database. However, the manufacturing library diversity is smaller than the amino acid sequence space, and the diversity analyzed exhaustively is smaller than manufacturing library diversity.

For this reason, we were not able to apply the high-potential protein to utilize in industry circles. Furthermore, the purpose function variants couldn’t conclude as part of the manufacturing library. The process of the library design including the purpose variants was required. In the directed evolution for antibody fragment, the mutagenesis processes established. In the case of affinity maturation, complementarity determining regions (CDRs) mutated to the degenerate codon which was NNK codon for all amino acids including a stop codon in all CDRs residues or the directional codon for partial amino acids from amino acids frequency in each residue. In the case of the enhancement of thermal stability, a few residues or the loop structure region in the framework region (FR) excluding CDR from the fragment of variable domains (Fv) mutated to the known amino acids or all amino acids. However, these mutagenesis processes were difficult to design the library including the purpose variants. In this study, we constructed the new directed evolution approach for antibody fragments, the Library design cycle, which approaches optimized variants in sequence space. The library design cycle iterated three processes, which are determining the mutation positions, identifying the positive amino acids, and screening the designed library.

Second section: The Library design cycle for mutating lysine residues in CDR

In this section, we verified the second process in the library design cycle, identifying the positive amino acids, utilized

site-directed mutagenesis on determining residues. After the evaluation of the site-directed mutated variants, we chose the positive amino acids and designed the library which had only the positive amino acids on each residue. We selected the purpose variants from the designed library. In this time, we tried to make bivalent single-chain Fv (diabody) from anti-mouse podoplanin antibody PMab-1 for antibody-drug conjugate (ADC). ADCs constructed antibodies and small molecular organic drugs. Besides, the way of conjugation with antibodies and drugs was the utilization of the specific amino acids, which were cysteine and lysine. In the case of antibody fragments, the utilization of cysteine residues was difficult because the disulfide bonds in each VH and VL were important interactions for structure formation. For this reason, the utilization of lysine residues, which were on the protein surface, was able to conjugate with the drug; however, the lysine residues in CDRs affected binding affinity by drug conjugation. For the verification of determining the mutagenesis residues, we tried to mutate the lysine residues in CDRs to the other amino acids with comparable binding affinity.

PMab-1 diabody had three amino acids (VH58, VH62, VL24). We mutated each residue by 22c-trick, which constituted three codons (NDT, VHG, TGG) and expressed 20 amino acids by 22 kinds of codons. Nextly, each variant was cultivated by 96 deep-well plates and evaluated expression level and binding affinity by ELISA used by culture supernatant. As a result, we selected some positive amino acids in each mutagenesis residue to consider the decrease of binding affinity due to mutation in CDR and the sequence space constructed from the positive amino acids. The sequence space, including only positive amino acids in each residue, was 32. Finally, we designed the degeneration codons for expressing only positive amino acids and evaluated the expression level and the binding affinity of each variant in the designed library. As a result, four kinds of variants were identified as the purpose variants which compared to the same binding affinity of PMab-1 diabody. The three variants in the top 3 were expressed by *E. coli* and purified by immobilized metal affinity chromatography, size exclusion chromatography, and FLAG-tag affinity chromatography, so that the acquired one variant was about 1000 times lower dissociation constant (K_d) than PMab-1 diabody but hold about ten times lower dissociation rate constant (k_{off}) than PMab-1 diabody. Furthermore, the variant had a specific delivery capacity to kidney by pharmacokinetic assay. In conclusion, we showed that the utilization of saturation mutagenesis in mutagenesis residues was an effective way of identifying the positive amino acids.

Third section: Library design cycle for generating cross active antibody

In this section, we examined the possibility of the application of random mutagenesis as determining the mutation positions, which was the first section of the library design cycle. We addressed the development of cross active antibody which recognized target protein derived from both animal and human. We used mouse anti-human podoplanin antibody, LPMab-16, which weakly recognized rat podoplanin, and expression vector of the single-chain Fv (scFv) was prepared from the sequences of humanized LPMab-16 Fv. For improvement of affinity to rat podoplanin, we first mutated scFv

with error-prone PCR to make a variant library and selected the positive clones from the library by means of the phage display method. In the comparison of primary structure between positive and wild-type clones, we determined the "mutation positions" which should be mutated to promote affinity. The "mutation positions" were VL27, VL27E, VL30, and VL50.

In the secondary step, saturation mutagenesis was independently conducted at each "mutation position" to analyze the affinity for human and rat podoplanin. From the results, candidate amino acids which induce no inactivation were selected at each position. Additionally, we selected wild type (WT) amino acid as candidate amino acids. The sequence space including candidate amino acids was 64kinds. Next, the small-sized library was generated by conducting simultaneous saturation mutagenesis with the candidate amino acids. As a result, the small-sized library was included 19 clones with higher affinity for rat podoplanin. The highest affinity for rat podoplanin clone with 27 times higher affinity for rat podoplanin was able to be identified without inactivation for the binding to human podoplanin. In conclusion, we showed that random mutagenesis was effective as "determining mutation position" and the combination with "determining mutation position" by random mutagenesis and "selected candidate amino acids" by saturation mutagenesis was able to downscale the sequence space to exploring diversity.

Forth section: Library design cycle for constructing the thermal stability enhancement process

In this section, we provided the iterative application of the library design cycle and the utilization of the bioinformatics for "determining the mutation positions" and "selected candidate amino acids". We addressed the enhancement of thermal stability for the low thermal stability antibody fragment, which was the diabody constructed from mouse anti-rat podoplanin antibody PMab-2. When acquiring antibodies by immunization approach, antibodies were mutated for affinity maturation by hyper somatic mutagenesis. Hyper somatic mutagenesis mutates not only CDR but also FR to the positive amino acids. We surmised that hyper somatic mutagenesis to the FR residues affected the debased thermal stability in the case of antibodies fragment. We attempted to mutate rare amino acid residues in the FR to enhance the thermal stability of the diabody. Utilizing the antibody database "abYsis" detected the mutated residues to calculate the frequency in each residue. Each residue in VH and VL of PMab-2 had calculated the frequency and consensus amino acid. Firstly, we analyzed the influence of mutation on rare residues on thermal stability and binding affinity. We designed 5 variants (ab5, ab10, ab20, ab30, ab35) mutated residues, which had up to the threshold frequency (5%, 10%, 20%, 30%, 35%), to consensus amino acids. Incremental mutations of the rare residues with low appearance frequency to those with higher frequency resulted in generating appropriate variants for enhancing the thermal stability; while the binding affinity for target irregularly varied. However, the mutations for the residues within 10% frequency decreased the thermal stability of diabody and those within 20% frequency inactivated the binding

affinity. Accordingly, the mutation by the bioinformatics succeeded in the enhancement of thermal stability while inactivated binding affinity. Secondly, to improve the thermal stability of PMab-2 diabodies without inactivation of binding affinity, we applied the iteration of the small-sized library design in the estimative range. We selected ab5 as template protein and determined the threshold frequency (10%, 12%, 15%, 20%, 25%, 30%) including within 7 residues. On the mutation for the residues within 5% frequency, we made the library where the wild-type or consensus amino acid appeared at the residues with the frequency of 5% to 10%, and the variants with comparable binding affinity to wild-type diabody were selected. The small-scale library was iterated to mutate the residues within 30%, and we found the variant which was selected from iterated library approach within 30% enhanced the thermal stability with the comparable binding affinity to wild-type. In conclusion, we showed that the bioinformatics utilization validated the application for “determining mutating position” and “selecting candidate amino acids”. Furthermore, the iteration of small-sized library design and selection was an effective approach for acquiring the purpose variant.

In this study, we constructed library design cycle, which improves the sequence space problem of directed evolution. In the future, we expected that proteins other than antibodies were modified function expeditiously by library design cycle.